GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: December 27, 2001, 16:50:10 ; Search time 28.15 seconds
(without alignments)
1823.860 Million cell updates/sec

Title: US-09-830-647-1
Perfect score: 3510
Sequence: 1 MNSGAMRIHSKGHFQGGIQV...........SDNLLTAFFSSPSTSTFTGF 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3124	89.0	601	2	T02633	hypothetical prote
2	187	5.3	1790	2	S67593	transport protein
3	178	5.1	1642	2	T08880	NMDA receptor-bind
4	173	4.9	1365	2	T30824	Imp1 protein - Myc
5	170.5	4.9	911	2	S51441	hypothetical prote
6	170	4.8	1516	2	E71619	hypothetical prote
7	166.5	4.7	1871	2	D96796	RAD2 endonuclease
8	165	4.7	1119	2	B70126	probable heat shoc
9	164.5	4.7	2401	2	T28676	surface-located me
10	164	4.7	646	2	F71620	rhopty protein - h
11	162	4.6	2269	2	T28677	hypothetical prote
12	161	4.6	1300	2	I53799	rhopty protein -
13	161	4.6	1356	2	S32763	CG1 protein - huma
14	159.5	4.5	1392	2	A44336	kinectin 1 - human
15	159.5	4.5	1427	2	S22695	microtubule-vesicl
16	158	4.5	980	2	E71606	restin - human
17	158	4.5	1271	2	A45555	hypothetical prote
18	157.5	4.5	1979	2	C71622	glutamate rich pro
19	157.5	4.5	3488	2	T34418	hypothetical prote
20	156.5	4.5	1690	2	T13030	microtubule bindin
21	155.5	4.4	1056	2	S55151	probable membrane
22	155.5	4.4	1577	2	T19722	hypothetical prote
23	155	4.4	1211	2	T27522	hypothetical prote
24	154.5	4.4	2663	1	S28261	centromere protein
25	154	4.4	1164	2	T24806	hypothetical prote
26	153.5	4.4	1017	2	PC4035	cell-cycle-depende
27	152.5	4.3	1115	2	T41342	probable coiled-co
28	152.5	4.3	1738	2	T14867	interaptin - slime
29	152.5	4.3	3259	1	A56539	giantin - human
30	152	4.3	1312	1	BWBYDL	RAD50 protein - ye
31	152	4.3	1980	2	S54307	myosin heavy chain
32	151.5	4.3	1005	2	A64465	hypothetical prote
33	151.5	4.3	1039	2	S62509	probable vesicular
34	151.5	4.3	1044	2	T50213	probable vesicular
35	151.5	4.3	2104	2	T38774	myosin-3 heavy cha
36	151.5	4.3	3147	2	T18674	hypothetical prote
37	150.5	4.3	1127	2	T28317	ORF MSV156 hypothe
38	150.5	4.3	1313	2	F96673	major merozoite su
39	150.5	4.3	1726	2	A45948	probable nuclear p
40	150.5	4.3	1837	2	T41023	hypothetical prote
41	149.5	4.3	1088	2	T18559	chromosome-associa
42	149.5	4.3	1231	2	S70553	major merozoite su
43	149.5	4.3	1726	1	SAZQGM	hypothetical prote
44	149	4.2	1650	1	T18444	hypothetical prote
45	149	4.2	2245	2	T18278	myosin heavy chain

ALIGNMENTS

RESULT 1
T02633
hypothetical protein RG135C18.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
C;Accession: T02633
R;Kellen, J.; Burkhart, J.
submitted to the EMBL Data Library, June 1998
A;Description: The sequence of Homo sapiens BAC clone RG135C18.
A;Reference number: Z14683
A;Accession: T02633
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-601 <KEL>
A;Cross-references: EMBL:AC005164; NID:g3242749; PIDN:AAC23786.1; PID:g3242750
A;Genetics:
A;Map position:
A;Introns: 60/3; 77/3; 101/1; 126/3; 139/1; 154/2; 197/2; 235/3; 277/2
A;Note: WUGSC:H_RG135C18.1

Query Match 89.0%; Score 3124; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.8e-166;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  74 RVEEFLSKDISYLISNKKEAKFAQTLGRISPVPSPSESAYTAETTSPHPSHDGSSFKSPDT 133
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 RVEEFLSKDISYLISNKKEAKFAQTLGRISPVPSPSESAYTAETTSPHPSHDGSSFKSPDT 60

Qy 134 VCLSRGKLLVEKAIKDIHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKKKELYLLKKSS 193
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 VCLSRGKLLVERAIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQKKKELYLLKKSS 120

Qy 194 TSVRDGGKRVGSGAQKTRTGRLKKPFVKVEDMSQLYRPFYLQLTNMPPINYSIQKPCSPF 253
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TSVRDGGKRVGSGAQKRTGRLKKPFVKVEDMSQLYRPFYLQLTNMPPINYSIQKPCSPF 180

Qy 254 DVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKKKRGYCECCLQKYEDLETHLLSE 313
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 DVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKKKRGYCECCLQKYEDLETHLLSE 240

Qy 314 QHRNFAQSNQYQVVDDIVSKLVFDFVEYEKDIPKKRIKYSVGSLSPVSASVLKKTEQKE 373
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 QHRNFAQSNQYQVVDDIVSKLVFDFVEYEKDIPKKRIKYSVGSLSPVSASVLKKTEQKE 300

Qy 374 KVELQHISQKDCQEDDTTVKEQNFLYKETQETEKKLLFISEPIPHPSNELRGLNEKMSNK 433
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 KVELQHISQKDCQEDDTTVKEQNFLYKETQETEKKLLFISEPIPHPSNELRGLNEKMSNK 360

Qy 434 CSMLSTAEDDIRQNFTQLPLHKNKQECILDISEHTLSENDLEELRVDHYKCNIQASVHVS 493
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CSMLSTAEDDIRQNFTQLPLHKNKQECILDISEHTLSENDLEELRVDHYKCNIQASVHVS 420
```

```
Qy   494  DFSTDNSGSQPKQKSDTVLFPAKDLKEKDLHSIFTHDSGLIITNNSQEHLTVQVQAKAPFHT  553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  DFSTDNSGSQPKQKSDTVLFPAKDLKEKDLHSIFTHDSGLIITNNSQEHLTVQVQAKAPFHT  480

Qy   554  PPEEPNECDFKNMDSLPSGKIHRKVKIILGRNRKENLEPNAEFDKRTEFITQEENRICSS  613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  PPEEPNECDFKNMDSLPSGKIHRKVKIILGRNRKENLEPNAEFDKRTEFITQEENRICSS  540

Qy   614  PVQSLLDLFQTSEEKSEFLGFTSYTEKSGICNVLDIWEEENSDNLLTAFFSSPSTSFTG  673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  PVQSLLDLFQTSEEKSEFLGFTSYTEKSGICNVLDIWEEENSDNLLTAFFSSPSTSFTG  600

Qy   674  F  674
          |
Db   601  F  601
```

RESULT 2
S67593

transport protein USO1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2555; protein YDL058w
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
C;Accession: S67593; A38455; S30782
R;Bloecker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67593
A;Accession: S67593
A;Molecule type: DNA
A;Residues: 1-1790 <BLO>
A;Cross-references: EMBL:Z74106; NID:g1431058; PID:e253003; MIPS:YDL058w
A;Experimental source: strain S288C
R;Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
J. Cell Biol. 113, 245-260, 1991
A;Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
A;Reference number: A38455
A;Accession: A38455
A;Molecule type: DNA
A;Residues: 1-389, 'TA',392-724,'S',726-1790 <NAK>
A;Cross-references: GB:X54378; NID:g4777; PIDN:CAA38253.1; PID:g4778
A;Note: the authors translated the codon ACT for residue 768 as Ile
R;Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993
A;Description: An integrin analogue in Saccharomyces cerevisiae.
A;Reference number: S30782
A;Accession: S30782
A;Molecule type: DNA
A;Residues: 71-846,'E',848-923,'K',925-1252,'I',1254-1318,'V',1320-1460,'S',1462-1580,'S'
A;Cross-references: EMBL:L03188
A;Gene: SGD:USO1; INT1
A;Genetics:
C;Genetics:
A;Cross-references: SGD:S0002216; MIPS:YDL058w
A;Keywords: coiled coil; transmembrane protein
A;Map position: 4L
F;326-342/Domain: transmembrane #status predicted <TM1>
F;394-410/Domain: transmembrane #status predicted <TM2>
F;611-633/Domain: transmembrane #status predicted <TM3>

```
Db  1084  REELEAELAA------------YKNLKELETKLETS----ERALKE---VKENE---------- 1119
                :|:|:|:         ||||:|:         ||:         :|
Qy   169  ILHIDDIRYYIEQK----KRELYLLKKSSTSVRDGGKRVGSGAQKTRTGRLKPFFVKVED    224
Db  1120  -EHLKEEKIQLKEATETKQQLNSLRANLESLEKEHEDLAAQLKKYEEQIANKERQYNEE   1178
           :|         :|| :          ::       :|         :
Qy   225  MSQLYRPFYLQLTNMFPINYSIQKPCSSPFDVD---KPSSMQKQTQVKLRIQTDGGDKYGGT  281
Db  1179  ISQLND----EITSTQQENESIKKKNDELEGEVVKAAMKSTSEEQSSNLK-KSEID------  1226
              :|||                  :  |||
Qy   282  SIQLQLKEKKKKGYC---------------------------ECCLQKYE--DLETHLLSEQH  315
                :| :|| ||                               ::|:|:    ::  ::|
Db  1227  ALNLQIKELKKKNETNEASLLESIKSVESETVKIKELQDECNFKEKEVSELEDKLKASED   1286
Qy   316  RNF-----AQSNQYQVVDDIVSKLIVEFDVFVEYEKDTPKKKRIKYSVGSLSPVSASVLKKTEQ  371
           :::|          ::        ||         ::|  :|        ||||
Db  1287  KNSKYLELQKESEKIKEELDAKTTELKIQLEKIINLSKAKERSESELSR-----LKKTSS   1341
Qy   372  KEKVELQHISQKDCQEDDTTVKEQNFLYKETQEFEKKLL-----FISEPIPHPSNELRGL   426
              :|:||        ::  | :          ||:|:::     ||    :  :  :|
Db  1342  EER----KNAEEQLELKNEIQIKNQAFEKERKLLNEGSSTITQEYSEKINTLEDE        1393
Qy   427  NEKMSNKCSMLSTAEDDIRONFTQLPLH----KNKQECILDISEHTLS-----------    470
              :|        :|  |:          ::|||         |:|
Db  1394  LIRLQNENELKAKEIDNTRSELEKVSLSNDELLEEKQNTIKSLQDEILSYKDKITRNDEK   1453
Qy   471  ------ENDLEELR-----VDHYKCNIQ--------------------------ASV    490
                 |  :|:|         ::|||         
Db  1454  LLSIERDNKRDLESIKEQLRAAQESKAKVEEGLKKLEEESSKEKAELEKSKEMMKKLEST  1513
Qy   491  HVSDPFSTDNSGSQPKQKSDTVLFPAKDLKEKDLHSIFTHDSGLIIT-INSSQE------H  542
           |  :||||||||||||||||||||||||||||||||||||||||||| |  |:      |
Db  1514  IESNETELKSSMETIRKSDEKLEQSKKSAEEDIKNLQHEKSDLISRINESEKDIEELKSK   1573
Qy   543  LTVQAK---APFHPEEPNECDFKNMDSLPSGKIHRKVKIILGRNRKENLEPNAEFDKRT    600
           |:||      | ||||||||||||||||||||||||||||||||||||||||||||||
Db  1574  LRIEAKSGSELETVKQELNNAQEKI------RINAEENTVL-KSKLEDIERELK-DKQA   1624
Qy   601  EF-ITQEENRICSSPVQSL---LDLFQTSEEKSE    630
           || |||||||||||||||     |||||||||
Db  1625  EIKSNQEEKELITSRLKELEQELDSTQQKAQKSE  1658
```

Query Match   5.3%; Score 187; DB 2; Length 1790;
Best Local Similarity 19.9%; Pred. No. 0.016;
Matches 150; Conservative 133; Mismatches 273; Indels 198; Gaps 31;

```
Qy     2  NSGAMRIHSKGHFQQGIQVNEKNRPSLKSLKTDNRPEKSKCKPLWGKVFYLDLLPSVTIS   61
          ::|:|:   ::    ::  ::   ::| |  ::
Db   978  NESLKAVEESKNESSIOLSNLONKIDSMSQEKEN------------FOIERGSIEKN    1023

Qy    62  -ERLQDIKDLGGRVEEFLSK----------DISYLISNKKEAKFA--QTLGRISPVPSP  108
           :||    |:  |:          ||||| |:      :|  :  |:
Db  1024  IEQLKKTISDLEQTKEEIISKSDSSKDEYESQISLIKEKLETATATANDENVNKISELTKT  1083

Qy   109  ESAYTAETTSPHSPHSDGSSFKSPDTVCLSRGKLLVEKAIKDHDFPIPSNSIISNALSWGVK  168
```

RESULT 3
T08880

NMDA receptor-binding protein yotiao - human
N;Alternate names: protein D2555
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T08880
R;Lin, J.W.; Wyszynski, M.; Madhavan, R.; Sealock, R.; Kim, J.U.; Sheng, M.
J. Neurosci. 18, 2017-2027, 1998
A;Title: Yotiao, a novel protein of neuromuscular junction and brain that interacts w
A;Reference number: T08880
A;Molecule type: mRNA
A;Accession: T08880
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Reference number: Z16511; MUID:98151389
A;Residues: 1-1642 <LIN>
A;Cross-references: EMBL:AF026245; NID:g2623067; PIDN:AAB86384.1; PID:g2623068
A;Genetics:
C;Genetics:
A;Map position: 7q21-22
C;Keywords: brain; cerebral cortex; coiled coil; neuromuscular junction; skeletal mus

Query Match   5.1%; Score 178; DB 2; Length 1642;
Best Local Similarity 21.1%; Pred. No. 0.044;
Matches 115; Conservative 104; Mismatches 184; Indels 142; Gaps 27;

```
Qy   179  IEQKKKELYLLKKSSTSVRDGGKRVGSGAQKTRTGRLKRPFVKVEDMSQLYRPFFYLQLTN  238
          ||:| ::|:||:|::|                    ::|
Db   351  IEEEKKTLELKDKLITADKILGELQOEIVQKNQ--EIKNMKLELTNSKOKEROSSEEIKQ   408

Qy   239  MFPINYSIQK----PCSPFDVDKPSSMQKQTQVK---LRIQTDGDKYGGTSIQL----    285
```

```
RESULT    4
T30822
lmp1 protein - Mycoplasma hominis
C;Species: Mycoplasma hominis
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 07-Dec-1999
C;Accession: T30822
R;Jensen, L.T.; Ladefoged, S.; Birkelund, S.; Christiansen, G.
Infect. Immun. 63, 3336-3347, 1995
A;Title: Selection of Mycoplasma hominis PG21 deletion mutants by cultivation in the pre
A;Reference number: 218884; MUID:95369882
A;Accession: T30822
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1365 <JEN>
A;Cross-references: EMBL:U21962; NID:g790243; PIDN:g790244; PIDN:AAA81013.1
C;Genetics:
A;Gene: lmp1
A;Genetic code: SGC3

Query Match    4.9%;  Score 173; DB 2; Length 1365;
Best Local Similarity  19.18%;  Pred. No. 0.066;
Matches  132;  Conservative 131;  Mismatches 257;  Indels 170;  Gaps 30;

Db   409 LMGTVEELQKRNIHKDSQFETDIVQRMEQETQRKLEQLRAELD-EMYGQQIVQMKQELIRQ 467
         |:||                                                  ||
Qy   286 ---QLKEKK--KKGYCECCLQKYEDLETH------------LLSEQHRNFAQSNQ--- 323

Db   468 HMAQMEEMKTRHKGEMENALRSYSNITVNEDQIKLMNVAINELNIKLQDTNSQKEKLKEE 527
Qy   324 ----YQVVDDIVSKLVF--DFVEYEKDTPKKRIKY----SVGSLSPVSASV 365

Db   528 LGLILEEKCALQROLEDLVEELSFSREQIQRARQTIAEQESKLINEAHKSLSTVEDLKAEI 587
Qy   366 LKKTEQKEKVELQHISQKDCQEDDTT----------VKEQNFL---YKETQETE----KK 408

Db   588 VSASESRKELELKH----------EAEVTNYKIKLEMLEKEKNAVLDRMAESQEAELERLRTQ 640
Qy   409 LLFISE-----------PIPHPSN---------ELRGLNEKMSNKCSMLSTAEDDI 444

Db   641 LLFSHEEELSKLIKEDLEIEHRINIEKLIKDNLGIHYKQQIDGLQNEMSQKIETMQFEKDNL 700
Qy   445 RQNFTQLPLH----KNKQECILD--ISEHTLSENDLE-ELRV----DHYKCNIQASVHVS 493

Db   701 ITKQNQLILEISKLKIDLLQQSLVNSKSEEMTLQINELQKEIEILRQEEKEKGTLEQEVQEL 760
Qy   494 DFSTDNGSGQPQKQSDTVLFPAKDLKEKDLHSTFTHDSGLIINS--SQEHLTVQAKAPF 551

Db   761 QLKTELLELEKQMEKEN-------DLQEK----FAQLEAENSILKDEKKTLEDMLKI 805
Qy   552 HTPPEEPNECDFKNMDSLPS------GKIHRKVKIILGRN---RKENLEPNAEFDK-RIEFI 603

Db   806 HTPVSQQERLIF--LDSIKSKSKRDSVWEKEIEILIEENEDLKQQCIQLNEEIEKQRNTFS 863
Qy   604 TQEEN 608

Db   864 FAEKN 868


RESULT    5
S51441
hypothetical protein YLR309c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L2142.5
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
C;Accession: S51441
R;Pauley, A.
submitted to the EMBL Data Library, November 1994
A;Description: The sequence of S. cerevisiae cosmid L2142.
A;Reference number: S51437
A;Accession: S51441
A;Molecule type: DNA
A;Residues: 1-911 <PAU>
A;Cross-references: EMBL:U17247; NID:g577216; PIDN:g577221; MIPS:YLR3
A;Cross-references: SGD:S0004300; MIPS:YLR309c
C;Genetics:
A;Gene: SGD:IMH1
A;Map position: 12R

Query Match    4.9%;  Score 170.5; DB 2; Length 911;
Best Local Similarity  19.3%;  Pred. No. 0.053;
Matches  137;  Conservative 125;  Mismatches 227;  Indels 221;  Gaps 33;

Db   376 NT---IQQAKTELEKEVQKADQAIRKSNNTASMQSAKSSLDAKVAE------- 417
Qy   233 YLQLTNMPFINYSIQKPCSPFDVDKPSSMQKQTQVKLRIQ-----TDGDKYGGTSIQLQ 286

Db   418 ----ITKKLFETFNKDKEAKFNELKQTRNQIQEFINTNKNNPNYSELISQLT 464
Qy   287 LKEKKKKGYCECCLQKYEDLETHLLSEQHRNFAQSNQYQVVDDIVSKLVFDFVEYEKDTP 346

Db   465 SKRDSKNSVTDS--SNKSDIES-ANTELKQALAKANADKVQADNLAKSI--- 510
Qy   347 KKKRIKRYSVGSLSPVSASVLKK--TEQKEKVELQHISQKDCQEDDTTVKEQNFLYKETQE 404

Db   511 -KEQLNNSVSNANTLSAKLTDKDNTIQQAKTEL----EKEVQKADQAIKSNN---- 560
Qy   405 TEKKLLFISEPIPHPSNELRGLNEKMSNKCSMLSTAEDDIRQNFTQLPLHKNKQECILDI 464

Db   561 MQSAKSSLDAKVAETTKKLETFNKDKEAKFNELKQTRNQIQEFINTNKNNPNYSELISQL 620
Qy   465 SEHTLSENDLEELRVDHYKCNIQASVHVSDPFSTDNS--------GSQPKQKSDTVLFPAK 516

Db   621 TSKRDSKNSVTD------SSNKSDIESANTELKQALAKANADKVQADNL---AK 665
Qy   517 DLKEKDLHSIFTHDSGLIINSSQEHLTVQAKAPFHTPPEEPNEC-DFKNMDSLPSGK-- 573

Db   666 SIKEQ-LNNSVSNANTLSAKLTDKDNTIQQAKTELEKEIOKANQAIKSNNTASMQSAKSS 724
Qy   574 IHRKVKIILGRNRKENLEPNAEFD--KRT-----EFITQEENRICSSPVQSLLDLFQTSE 626

Db   725 LDAKVAEITKKLETFNKDKEAKFNELKQTRNQIQEFINTKN---NPNYSELISQLTSK 780
Qy   627 EKSEFLGFTSYTEKSGICNYLLDIWEEENSD 656

Db   781 RDSK---NSVTDSS---NKSDI-ESANTE 802
```

RESULT 6

```
E71619
RAD2 endonuclease PFB0265c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C;Accession: E71619
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:9901743
A;Accession: E71619 <GAR>
A;Cross-references: GB:AE001383; GB:AE001362; NID:g3845135; PIDN:AAC71842.1; PID:g3845135
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0265c

Query Match            4.8%;  Score 170;  DB 2;  Length 1516;
Best Local Similarity 19.5%;  Pred. No. 0.11;
Matches 138;  Conservative 115;  Mismatches 257;  Indels 198;  Gaps 31;

Qy    201  KRVGSGAQKTRT-GRLKPFFVKVEDMSQLYRPFYLQLTNMPFINYSIQKPCSPFDVDKPS  259
                |:|
Db    287  RR------KRNRNKGGITTGDISE------EETVDNSI                        317

Qy    260  SMQKTQVKLRIQTDGDKYGGTSIQLQLKEKKKKGYCECCLQKYEDLETHL-----       310
                ::|
Db    318  NTEEYDKLKENLQ--------ELQERYKD---CEDWKQKYEDIEAELKDAKELENS       362

Qy    311  ------LSEQHRNFAQSN-QYQVVDDIVSKLVFDFVEYE---------KDTP          346
                 |:|:|
Db    363  QLEKSAKELETLNTELIDTKKSLKEKNSELEEVRDMLRTVGNELVDAKDEIKESSSKQNE   422

Qy    347  KKKRIKYSVGSLSPVSASVL-----KKTEQKEKVEL-----QHISQKDCQEDDITVKEQN  396
                |:|
Db    423  EVKTVKLELDLRHKNATMIEAYEAKNTELRSKIELLSKKVEHLKNL-CTE---------   472

Qy    397  FLYKETQETEKKLLFISEPIPHPSNELRGLNEKMSNKCSMLSTAEDDIRQNFTQL-PLHK  455
                |:|:|
Db    473  ------KEKEQTTSQ------NKVAKLNEEI------SQLTYEKSNITKELTSLRTSYK   513

Qy    456  NKQECILDISEHT------------LSENDLEELRVDHYKCNIQ----ASVHVSDSFSTDN 499
                :|:|
Db    514  QKEKTVSYLEEQVKQFSEQKDVAEKSTEQLRDHAKISNRLDLLKKENETLLHNDIAKNSN  573

Qy    500  SGSQPKQKSDTVLPPAKDLKEK---DLHSIFTHDSGLI-TINSSQEHLTVQAKAPFHPPE  556
                 |:|
Db    574  SVEEYLKENGKLSERLNILQEKVNTLQNVKSNSNEHIDSIKRQCEELNVKLKE------   626

Qy    557  EPNECDEKNMDSLPSGKIHRKVKIILGRNRKENL----------EPNAEFDKRTEFI     603
                |:|:|
Db    627  ------STKKILSL-EDELNEYANIVQDKTREANTLRRLVSDSDSQTDDSSKQRELENKLAYL 680

Qy    604  TQEENRICSSPVQSLLDLFQTSEEKSEFLGFT-SYTEKSGICNVLDIWEE            652
                |:|:|
Db    681  TDEKNKL------EAELDL-QTSRKATELQEKWHTVTELKSEIHALKLRREE          724
```

RESULT 7

```
D96796
E71619
probable heat shock protein, 53413-59028 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96796
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
ansen, N.F.; Hughes, B.; Hulzar, L.;
R;Theologis, A.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: D96796
A;Molecule type: DNA
A;Residues: 1-1871 <STO>
A;Cross-references: GB:AE005173; NID:g6143906; PIDN:AAF04452.1; GSPDB:GN00141
C;Genetics:
A;Gene: F28016.15
A;Map position: 1
A;Status: preliminary

Query Match            4.7%;  Score 166.5;  DB 2;  Length 1871;
Best Local Similarity 20.2%;  Pred. No. 0.23;
Matches 148;  Conservative 109;  Mismatches 267;  Indels 207;  Gaps 33;

Qy    200  GRKVGSGAQKTRTGRLKKPFFVKVEDMSQLYRPFYLQLTNMPFINYSIQKPCSPFDVDKPS 259
                :|:|
Db    790  ------KKNIYNKNNNDSN--KTFFLKIEN-----                            810

Qy    260  SMQKTQVKLRIQTDGDKYGGTSIQLQLKEKKKKGYCECCLQK------YEDLETHLLS     312
                |:|
Db    811  ------EFKKDLLLDDSQIFGDSLLADIKE--YNYTADNLDNNNENKSLYEDGENFIT-    860

Qy    313  EQHRNFAQSNQYQVVDDIVSKLVFDFVEYE-------KDTPKKKRIKYSVGSLSPVSASVL 366
                |:|:|
Db    861  --RNEPITNEYEEKNNII--YISDEQKYNEEDIIFKDKIKEKEKNNDTSSDDFENCSVQ   915

Qy    367  KKTEQKEKVELQHISQKDCQEDDITVKEQNFLYKETQETE--KLLFISEPIPHPSNEL     423
                 |:|
Db    916  EKIYVNEKIEEYNNKNDDKSSSSSSIILEEIKYKKEKKDELVSPNLCVLLDEFEH-SNDL  974

Qy    424  RGLNEKMSNKCSMLSTAEDDIRQNFTQLPLHKNKQECILDISEHTLS---ENDLEELRVD  480
                |:|:|
Db    975  EN---------NYISVSSDDMKTN----------VSKNNITGVKENKVDKTNVE       1009

Qy    481  HYKCNIQASVHVSDFSTDNSGSQPKQKSDTVLPPAKDLKEKDLHSIFTHDSGLIINS--   538
                :|:|
Db   1010  YDKKGDDGVIEIS-FEDSHKLEESKFDDNNIYDNDDELEKNLSKDYISDVDKNHVNIIY   1068

Qy    539  ----SQEHLTVQAKAPFHPPEEPNE--C------DFKNMDSLPSGKIH             575
                 |:|
Db   1069  NIERGEDERENEFVENKIQ-STESHKSNEFICTENKSLRKQYMSKEDISNVRILKSDDIN 1127

Qy    576  RKVK-----IILGR----NRKENLEPNAE-----FDKRTEFITQEENRICSSPVQSLL   619
                |:|:|
Db   1128  NLSKQNYFEILLDKKQVMDNFQMNIEQCNNDKLKEDKLDEGAYFEYLEDNKIIDSYIK--- 1184

Qy    620  DLFQTSEEKSEFLGFTSYTEKSGICNVLDIWEEENSD-NLLTAFFSSP            666
                |:|:|
Db   1185  ---EINKENEELIKEYKKLKK---NNIEINDEMNDDIKLLNFFGIP              1225
```

```
Query Match            4.8%;  Score 170;  DB 2;  Length 1871;
Best Local Similarity 20.8%;  Pred. No. 0.11;
Matches 138;  Conservative 115;  Mismatches 257;  Indels 198;  Gaps 31;

Qy     29  LKSLKTDNRPEKSKCKPLWGKVFYLDLPSVIISEKLQK------DIKDLGGRVEEFLSKD  82

Qy     83  ISTL---ISNKEAKFAQTLGRISPVPSPESAYTAETTSPHPSHDGSSSFKSSPTVCLSRG 139
Db    646  VRMIKTDFEIEKLRIK-----MAENIQSVEGELLINKDLSKNTDNINIKD           690

Qy    140  KLLVEKAIKDHDFIPSNSILSNALSWGVKILHHDDIRYYIEQKKELLYLLLKSSTSVRDG 199
Db    691  YNYLQKKKSKKKKFLNDI------LNTYFTESKYQDLYKVKGEESKEDI-KNQI        740

Db    741  DFVTQECYRNNDIIRDTHDKSDIF----KNIKIDN----NKKYEIIYNLELEQEEINE-  789
```

```
Query Match            4.7%;
Best Local Similarity
Matches 148;  Conservative 109;

Qy     17  GIQVKNEKNRPSLKSLKTDNRPEKSKCKPLW---GKVFYLDLPSVISEK-----LQKDI   68
                :|:|
Db    907  GIE-EHELHEPKIHKERDNNRVTGAK-EPSGQEKGEKEEKIVRSMTITENDNSIDVQETK  964
```

```
Qy   69  KDLGRVEEFLSKDISYLI----------------SNKKE---------AKFAQTLGRISPVPSPESA 111
              |:|:||            |:|||           |:|:||
Db  965  KERPGRLE---SHDKRYKIQELLMEAGHNDRKEEEQNENVTAEVELETERVSSKKVOEGK 1021

Qy  112  YTAETTSP-HPSHDGSSEFKSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKIL 170
Db 1022  MEDDNSGKFHEFEERKSYEDWTHEKRKRKVLVE----EEETYPRDKHTCG------- 1068

Qy  171  HIDDIRYYIEQKKKELYLLKKSSTSVRDGGKRVGSGAQKTRTGRLKKPFVKVEDMSOLYR 230
Db 1069  --EDHNDHKEEEQKENVIAKAELNTEEDSPFKKV-EEIEKQDHGELKRSMVQA------- 1117

Qy  231  PFYLQLTNMPFINYSIQKPCSPFDVDKPSSMQKQTQVKLRIQT-DGD------------KYG 279
              |: ||          ||||          |||
Db 1118  --------------KRQETEEKDKTRAMEKNETVERRKQTKDGSLGKIREGEDPELG 1160

Qy  280  G----------TSIQLQLKEKKKGYCECCLQKYEDLETHLLSEOHRNFAQSNQY 324
Db 1161  GHERRGEEDRIEELVETEISDHKEKVKKD--EDYILRSQDTGKVDLGERERRSKORKIH 1218

Qy  325  QVVDIVSKLIVFDEVEYEKDTPKKRRI--KYSVGSLSPVSASVLKKTEQKEKVELQHISQ 382
Db 1219  KSVEDEIGD--------QEDEAEEAAAVSRNENGSSRKVQTIEEESEKHKEQNKIPETSN 1272

Qy  383  KDCGOEDDTVVKEQNFLYKETQETERKLLFISEPIPHPSNELRGLNE---------KMSNK 433
Db 1273  PEVNED--------EERVEKEFKEVEAHV---------------QELEGKYTENCKQDDGEGRREER 1317

Qy  434  CSMLSTAEDDIRONFTQLPLHKNKOECILDISEHTLSENDLEELRVDHYKCNIQASVHVS 493
Db 1318  GKOGMTAENMLRORFKT---KSDDGIVRKIQETREEEPD------EKKSQESSSHVV 1365

Qy  494  DFSTDN----SGSOPKQKSDTV--LFPAKDLKEKDLHSIFTHDSGLITINSSQEHLTVQA 547
Db 1366  KIVAEDGSLRNGLEFSEKESTVSKMIKLDESKEKEEHK---------------------- 1403

Qy  548  KAPFHTPPEPNEC---------DFKNMDSLPSGKIHRKVKIILGRNRK-ENLEP------NA 594
Db 1404  -KIRKPTEERSNAPVIEKQGNKKNAEEEMQDKIDRR----GKNQEIKGQEPYGVLRNG 1456

Qy  595  EFDKRTEFITQEE---------NRICSSPVQSLLDLFQTSEEKSEFLGFTSYTEKSGICNVLDI- 649
Db 1457  EHDKITEYHRGEEKGTAENVSSTKIQQTKDELEKPRKPSEI-----SENHIHEFMDSS 1510

Qy  650  ----WEENSD 656
              |: ||
Db 1511  QSQDIEEKGSD 1521
```

RESULT   8

B70126
surface-located membrane protein 1 (lmp1) homolog - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 11-Jan-2000
C;Accession: B70126
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
  son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
  ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
  Nature 390, 580-586, 1997
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Authors: Smith, H.O.; Venter, J.C.
A;Reference number: A70100; MUID:98065943
A;Accession: B70126
A;Molecule type: DNA
A;Residues: 1-1119 <KLE>
A;Cross-references: GB:AE001131; NID:g2688098; PIDN:AAC66595.1; PID:g2688810
A;Experimental source: strain B31
A;Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat ho
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
F;742-774/Domain: tetratricopeptide repeat homology <TT2>
F;775-808/Domain: tetratricopeptide repeat homology <TT3>
F;809-842/Domain: tetratricopeptide repeat homology <TT4>
F;843-876/Domain: tetratricopeptide repeat homology <TT5>
F;911-943/Domain: tetratricopeptide repeat homology <TT6>
F;944-977/Domain: tetratricopeptide repeat homology <TT7>
F;979-1010/Domain: tetratricopeptide repeat homology #status atypical <TT7>
F;1011-1044/Domain: tetratricopeptide repeat homology <TT8>

Query Match   4.7%; Score 165; DB 2; Length 1119;
Best Local Similarity   20.6%;
Matches 161; Conservative 133; Mismatches 289; Indels 200; Gaps 44;

```
Qy   22  NEKNRPSLKSLKTDNRPEKSKCKP----LWGKVF--YLDLPSVVISEKLOKDIKDLG----72
              |:||: :|||: :::|:|               :||:   :||| :|||
Db  249  NNNNTTSLKKISSNSQKESELSPPSQTIIGKIYRPY---SYLIKKELYEILDDINTGRV 304

Qy   73  -----GRVEEFLSKDIS--------YLISNKKEAKFAQTL----GRISP--VPSPESAY- 112
Db  305  --TAETTSPHPSHD-GSSEFKSPDTVCLSRGKLLVEKAIKD-HDFI--PSNSILSNA 162

Qy  113  ----TAETTSPHPSHD-GSSEFKSPDTVCLSRGKLLVEKAIKD-HDFI--PSNSILSNA 162
Db  305  TLGKNRLKELIKKGLSNKFQKVNELIENSKNKEASNLLITLIKKDIEPNLINIPKDPYKK 364

Qy  163  LSWGVKILHIDDIRYYIEQ-KKKELYLLKKSSTSVRDGG---KRVGSGAQKTRTGRLKKP 218
Db  365  EIFQLDKIEDKKPQYLEDLKSKVHSIKPIDLENTKSR-QQAIKDLNEFLKNNPNDAQASKT 423

Qy  219  FVKVEDM-SQLYRFFYLQLTNMPFINYSIQKPCSPFDVDKPSSMQ--KQTQVKLRIQTDG 275
Db  424  LAQANKIQHLEDLKSKVHSIKPIDLENTKSRQQAI-KDLNEFLKNNPNDAQASKTLAQANK 483

Qy  276  D-KYGGTSIQ-LQLKEKKKKGYCECCLQKYEDLETHLL----SEQHRNFAQSNQYQVVD 328
Db  484  IQHLEDLKSKVHSIKPIDLENTKSRQQAI-KDLNEFLKNNPNDAQASKTLAQANKIQHLE 542

Qy  329  DIVSKL-----------VFDFVEYEKDTPKKKR----------------------IKYSV 355
Db  543  DLKSKVHSIKPIDLENTKSR------QQAIKDLNEFLKNNPNDAQASKTLAQAYENNGDLL 596

Qy  356  GSLSPVSASVLKKTEQ--KEKVELQHISQKDCQEDDTVVKEQNFLYKETQETEKKLLFTS 413
Db  597  DLKSKVHSIKPIDLENTKSRQQAIKDLNEFLKNNPNDAQASKTLAQANKIQHLEDLKSRV 656

Qy  414  EPIPHPSNE----LRGLNEKMSNKCSMLSTAEDDIRONFTQLPLHKNKOECILDISEHT 468
Db  657  HSIKPIDLENTKSRQQAIRDLNEFLKNNPNDAQASKTLAQANKIQHLE--DLKSKVHSI- 713

Qy  469  LSENDLEEL--------RVDHYKCNI--------QASVHVSDFSTDNSGSQPKQKSDTVLFP 514
Db  714  KPIDLENTKSRQQAIKDLNEFLKNN------PNDAQASKTLAQAYENNGDLL------- 759

Qy  515  AKDLKEKDLHSIFTHDSG--LITINSSQEHL-----TVQAKAPFHTP--------PEEPNE- 560
Db  760  KAENAYEKIIKLINTQEDHYKLGIIRFKLKKYEHSIESFD---------- 799

Qy  561  ------CDFK---NMDSLPSGKIHRKVKII---LGRNRKENLEP---------NAEFDKRTEF-- 602
Db  800  -QTIKLDPKHKALHNKGIALMMLNKKRAIESFEKAIQIDKNYGTAYYQKGTAEEKNGD 858

Qy  603  ---ITQEENRICSSPVQSLLDLFQTSEEKSEFLGFTS--YTEKSGICNYLDIWE---EEN 654
Db  859  MQQAFASRKNAYNLDKNENYAL--KAGIVSNNLG-NFKQSEEYLNFFNANAKKPNEIAIY 915

Qy  655  --WEENSD 657
         | ::|
Db  916  NLSTAKFENNKLEESLEITNKAIDLNPEKSEYLYLKASINLKKENYQNAISLYSLVIEKN 975

Qy  655  SDN 657
         |:|
Db  976  PEN 978
```

RESULT   9

T28676
rhoptry protein - Plasmodium yoelii (fragment)
C;Species: Plasmodium yoelii
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C;Accession: T28676; A45521

R;Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332 1996
A;Title: Comparison of two members of a multigene family coding for high-molecular mass
A;Reference number: 2
A;Accession: T28676
A;Molecule type: protein
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-2401 <SIN>
A;Cross-references: EMBL:U36927; NID:g1041784; PID:g1041785; PIDN:AAB41263.1

R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A;Reference number: A45521; MUID:91101660
A;Accession: A45521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2260-2401 <KEE>
A;Cross-references: GB:M34281

Query Match
Best Local Similarity 20.6%;
Matches 155; Conservative 98; Mismatches 305; Indels 193; Gaps 34;
Score 164.5; DB 2; Length 2401;

```
Qy   11 KGHFQGGIQVKNEKNRPSLKSLKTDNRPEKSKCKPLWGKVFYLDLLPSVTISEKLQKDIKD 70
            :: :       ::|:  |   | ||   | |    |||
Db  421 KKHYNDQINIDNIKE----KEAKQNYDQFKEHMKTIPPNEMKYQKPSIEI--KIMKD--- 471

Qy   71 LGGRVEEFLSK-----------KFAQTLGRISPVSPSESAYTAETTSPHPS 122
          |   ||||||
Db  421 ----EFLSKYNKYNDFDKVYKEKVESEHNKFTELITNKIRTEVSDEEIKKYE---NKF 521

Qy  123 HDGSS--------------FKSPD----TVCLSRGKLLVEKAIKDHDFIPSNSILSN 161
          ::|
Db  522 NDSKSLINETKKSIEEEYQNINTLKKVDDYIKVCLNTNELITNCHNK------QTTLKD 574

Qy  162 ALSNGVKI----------LHIDDIRYIEQKKKELYLL------------KKSSTSVRDG 199
          |:  |:
Db  575 KLNQNIKTIKEINSIDKIYDKFENILTDKKTELETKFTGLSLNNHESNNKELLTYFYDL 634

Qy  200 GKRVGSGAQKTRGLKKPFVK-------VEDMSQLYRPFYLQLTNMPFINYSIQRPCSPFDV 255
          |:|                        |||
Db  635 KANLG------KNKENMLYKQENEKEKAVEDIKKK------------NVDINIKVSNIEI 676

Qy  256 DKPSSMQKQTQVKLRIQTDGDKRYGGTSIQLQLKEKKKKGYCECCLQKYEDLEHLLSEQH 315
          : :
Db  677 TIYTSI--------YNINEDTENEIGKSIEL------------LNTKVLEKVK 709

Qy  316 RNFAQSNQYQVVDDIVSKLV-FDEYEYEKDT----PKKKRIKYSVGSLS-PVSASVLKKT 369
          :  :     ::
Db  710 ANVTNLN------EIKEKLKDYDFQDFGKEKNIKYPDENKIKNDIDTLNQKIDKSIETLT 763

Qy  370 EQKEKVELQHISQKDCQEDD-TTVKEQNFLYKETQETEKKLLFISEPIPHPSNELRGLNE 428
          ::  ::
Db  764 EIKKNSE-NHIDEIKGQIDKLKKYPNKTMFNEDPKEIEKIENIVEKIDKKNIYKEI-D 821

Qy  429 KMSNKCSMLSTAEDDIRQNFTQLPLHKNKQEICLDISEHTLSENDLEELRVDHYKCNIQA 488
          ::|
Db  822 KLLNEISKIENDKTSL-EKLKNINLSYGK------SLGNLFLQQIDEEKKKAEH----TIKA 872

Qy  489 -SVHVSDFSTDNSGSQPKQKSDTVLFPAKDLKEKDLHSI-FTHDSGLITINSSQEH---- 542
          :  :
Db  873 MEAYIDDLDNIKKKSQEIEKEMNINMDIKMDIHKEMRALNISHDDYKIYHTTSKNHEEKI 932

Qy  543 ------------LTVQAKAPFHTPPEEPNEDCDFKNMDS-------LPSGKI 574
                       |:
Db  933 SDIRKNSLKIIQDFSEESYINDIKKELENVLESQNNNTDINQYLSKIENIYNILKLNKI 992

Qy  575 HR---RVKILGRNRKENLEPNAEFDKRTEFITQ-EEN---RICSSPVQSLLDLFQTSEE 627
          |||
Db  993 KKIIDKVKEYTDEIEKNNKKINAELSNSEKIITQLKENSSLKECQOSKIKSTIDDNYVSEC 1052

Qy  628 KSEFLGFTSY--TEKSGICNVLDIWEEENSD 656
```

Db 1053 IKNITNLKTYIYINEKKNINTYFKNAEEYNQN 1083

RESULT 10
F71620
hypothetical protein PFB0230c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
R;Gardner, M.J.
C;Accession: F71620
A;Accession: F71620
.;Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
R;Reference number: F71620
.;Reference number: A71600; MUID:99021743
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-646 <GAR>
A;Cross-references: GB:AE001381; GB:AE001362; NID:g3845124; NID:g3845124; PIDN:AAC71835.1; PID:g384
A;Cross-references: GB:AE001381;
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0230c

Query Match
Best Local Similarity 18.5%;
Matches 136; Conservative 116; Mismatches 229; Indels 256; Gaps 28;
Score 164; DB 2; Length 646;

```
Qy   39 EKSKCKPLWGKVFYLDLLPSVTISEKLQKDIKDLGGRVEEFLSKDISYLISNKKEAKFAQT 98
          :|
Db   10 EQNKMNTLHIKNFIMENLKVTEEIKHDKDINNLMRIEH--EEIKELISSNGKRYFME- 65

Qy   99 LGRISPVSPSESAYTAETTSPHPSHD---GSSFKSPDTV-----CLSRGKLLVEKAIKD 149
          :::
Db   66 --IRKIYFLMKKFHKEGYFPSSNKDVLKKQSFKRNKNIKNLQESIKKKNIQIOKLLKQ 122

Qy  150 H--------------DFIPSNSILSNALSWG------------ 166
          :
Db  123 YIILKGYYKNVCKKYRQENELLKSFSFSNNQSYTLNLKYSPHSRNRIFYPYTKLLR 182

Qy  167 ---VKILHIDDIRYYI----EQKKKELYLLKK--------------- 191
             ||:
Db  183 RKRLRRISHFKEDRVIHKGPLTKKKKKKIYINKKYIYIYIYIYIFFMFYSFIFIE 242

Qy  192 -SSTSVRDGGKRVGSGAQKTRGLKKPFVKVEDMSQLYRPFYL-----QLTNMPFINY 244
          ::|
Db  243 YFSNSI----------FRKYTHHKKRYKEIIQDILNDKLLINLFKRY 280

Qy  245 SIQKPCSPFDVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKKKKGYCECCLQKYE 304
          |:
Db  281 K----------EKYKKKKKLIHISSKKRK-DKRNLDLYCKKKKK----E 315

Qy  305 DLETHLL--SEQHRNFAQSNQYQVV----DDIVSKLVFDFVEYEKDTPKKRIKYSVGSL 358
          : :
Db  316 IIYTHLFLPTRLREKINKSSNYNYLNKEGENTINKEENILHKEE----- 361

Qy  359 SPVSASVLKKTEQK-EKVELQHISQKDCQEDDTTVKEQNFLYKETQETEKKLLFISEPIP 417
          ::
Db  362 -HLHKDEENYMKEEEENILHKDEEENILYKEEENILHKDEE---------- 403

Qy  418 HPSNELRGLNEKMSNKCSMLSTAEDDIRQNFTQLPLHKNKQEICLDISEHTLSENDLEEL 477
          :||
Db  404 ----------NILYKEEENI----LHKDEEENILHKEEENILHKDEEEN 438

Qy  478 RVDHYKCNIQASVHVSDFSTDNSGSQPKQKSDTVLFPAKDLKEKDLHSIFTHDSGLITIN 537
          : :
Db  439 ILYKEEENILHKEEANIIETKN--AEVKKKKNTL----RKKKKKEKKKNFLNDHMKEVTKN 492

Qy  538 SSQEHLTVQAKAPFHTPPEEPNECDFKNMDSLPSGKIHRKVKIILGRNRKENL------- 590
          :::
Db  493 DDDD------------DDDDDDDENNMI-----KVEEKQK-YNDEDGKENVSIDNVEE 532
```

```
RESULT 11
T28677

rhoptry protein - Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C;Accession: T28677; C45521
R;Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A;Title: A gene coding for a high molecular mass rhoptry protein of plasmodium yoelii.
A;Reference number: Z20508; MUID:95021522
A;Accession: T28677
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2269 <KEE>
A;Cross-references: EMBL:L27838; NID:g457145; PIDN:AAA21304.1
A;Cross-references: EMBL:L27838; NID:g457146; PIDN:AAA21304.1
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A;Reference number: A45521; MUID:91101660
A;Status: preliminary
A;Accession: C45521
A;Molecule type: DNA
A;Residues: 2131-2269 <KE2>
A;Cross-references: GB:M34283

Query Match    4.6%;  Score 162; DB 2;  Length 2269;
Best Local Similarity  17.4%;  Pred. No. 0.53;
Matches  149;  Conservative  152;  Mismatches  289;  Indels  266;  Gaps  34;

Qy      9 HSKGHFPQGGIQVKNEKNRPSLKSLKTDNRPEKSK-------------------------CKPLWGKVF 51
Db   1085 NEKIHIEDEV-ANNDKSNKAITSTKVSVEPFKTKIKINEIRTKSDDCLKETNDLEKQIS 1143

Qy     52 YLDLPS------------VTISE----KLQKDIKDLGGRVEEFLS---KDISYLISNKK 91
Db   1144 NLSIDTQETKLIENGKQLKTLEELLESLKKQKKNIEDQKKELDEVNSKIKINIENTVNQHK 1203

Qy     92 --EAKFAQTLGRISPVSPSPESAVTAEFTSPHPSHDGSSFKSPDTVCL----SRGKLLVE 144
Db   1204 KNYEIGIVEKINEIAKTNKNQIESTKELIKPTIQHIISPNANDLEGIDSDENLGKYNTE 1263

Qy    145 KAIKDHDFIPSNSILSNALSWGVK--------------------ILHIDIRY------YI 179
Db   1264 MGNIYEEFIKSYNLITNYLETVSKESITYNQIQNKRIDTQKELLKNIENVNKAKSYLDYI 1323

Qy    180 EQK--------KRELYLL-----------KKSSTSVRDGGKRVGSGAQKIRTGRLKKPFVKVEDMS 226
Db   1324 KENEFDRIVTHFKKLNIVNDNFKNEYSKVNEGFDNISNSINTVKNSTDENSLNILNQT 1383

Qy    227 QLYRPFYLQLTNMPFINYSIQ-----------------------KPCSPFDVDKP------ 258
Db   1384 --KEMYANIVNNTYYSYKYEAENIFRNIPKLANTLNIKIKNSSGIDLSKDIKIAILSYL 1440

Qy    259 -----------SSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKKKKGYCECCLQKYEDLE 307
Db   1441 DSKKEDTLIFIPSPQKKTETYTKI---SDSY---SILLDILKKSQE----------- 1488

Qy    308 THLLSEQHRNFA---------------------------QSNQ------YQVVDD 329
Db   1489 LKLIFENRLYEKVQATNELRGTILSDLKYKKEKILSEVKLLLHKSNELNKLSCNFQNYDT 1548

Qy    330 IVSKLVFDPVEYEKDTPKKKR-----IKYSVGSLSPVSASVLKKTEQKE----------- 373
Db   1549 ILESSKYDQVREKSNNYKQEREKLGIGIDFNVYDMEERKFNNDIKVIEELENNYDSSEENNI 1608

Qy    374 ---KVELQHISQK---DCQEDDTYVKEQNFLYKETQETEKKLLFISEPIPHPSNELRGLN 427
Db   1609 LQSKQRLKELTNKFNAEIKIKIDDKIIEKNDLLDKLIETRKNCMLFPH---------TTLAETLK 1663

Qy    428 EKMSNKCSMLSTAEDIRQNFTQLPLHKNKQECILDISEHTLSEND--LEELRVDHYKC 484
Db   1664 IKITDYSKFIESAT------------------KFSKEFLKYIGDTSNSLNDDIATLQLKYDLHQI 1710

Qy    485 NIQASVHVSDFSTDNSGSQOPKQKSDTVLFPAKDLKEKDLHSIFTHDSGLITINS--SQEH 542
Db   1711 NKYVISKLSDATNDNNNLIEKEKEATQAI---------KNLTKLFTIDSNNIDANALHNKI 1763

Qy    543 LTVQAKAPFHTPPEEPNECDFRNMDSLPSGKIHRKVKII----LGRNRKENLEPNAEFDK 598
Db   1764 QMVYFNSELH---------KSIESIK--QLYKKMIVFKLLNIGQINGKYFDISKQFDN 1810

Qy    599 RTEFITQEENRICSSPVQSLLDLFQ-TSEEKSEFL-------K 640
Db   1811 ----ILQLQESELTANLNDLKEIGQKISDKRKNFLHALNETPIPNFNTLKEIYHDIVYK 1866

Qy    641 SGICNVLDIWEEENSD 656
Db   1867 RQIDEINITSEENEN 1882


RESULT 12
I53799

CG1 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C;Accession: I53799
A;Molecule type: mRNA
A;Residues: 1-1300 <RES>
A;Cross-references: GB:L25616; NID:g409465; PIDN:AAB65853.1; PID:g409466
A;Status: preliminary; translated from GB/EMBL/DDBJ
R;Print, C.G.; Leung, E.; Harrison, J.E.; Watson, J.D.; Krissansen, G.W.
Gene 144, 221-228, 1994
A;Title: Cloning of a gene encoding a human leukocyte protein characterised by extens
A;Reference number: I53799; MUID:94314220
A;Genetics:
A;Gene: GDB:CG1
A;Cross-references: GDB:450341

Query Match    4.6%;  Score 161; DB 2;  Length 1300;
Best Local Similarity  20.2%;  Pred. No. 0.29;
Matches  147;  Conservative  130;  Mismatches  279;  Indels  172;  Gaps  34;

Qy     17 GIQVKNEKNRPSLKSLKTDNRPEKSKCKPLWGKVFVFLDLPSVTVISERLQKDIKDL---G 72
Db    262 GIQ-------KSGTKKLKTETDKENAEVK----FKDFLLSLKTMMFSEDEALCVVDLLKEKS 312

Qy     73 GRVEEFLSK-----DISYLISNKKEAKFAQTLGRISPVSPSPESAVTAEFTSPHPSHDGSSF 128
Db    313 GVIQDALKKSSKGELTTLIHQLQEKD--KLLAAVK-----EDAAATKDRCKQLTQEMMTE 365

Qy    129 KSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYYIEQRKKELYL 188
Db    366 KERSNVVITRMKDRIGTLEKEHNVFQNKRIHVSYQETQOMQ-MKFQQVR---EQMEAETAH 421

Qy    189 LKKSSTSVRDGGKRVGSGAQKTRTGRLKKPFVKVEDMSQLYRPFYLQLTNMPFINYSIQK 248
Db    422 LKQENGILRDAVSNTTNQLESKQSAELNK-----LRQDYARLVNELTEK---TGKLQQ 471

Qy    249 PCSPFDVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKKKKGYCECCLQKYEDLET 308
Db    472 E---EVQKKNAEQAATQLKVQLQEAERRWE---EVQSYIRKRTAEHEAAQQ---DLQS 520

Qy    309 HLLSEQHRNFAQSNQYQVVDDIVSK-----LVFDFVEYE-KDTPKKRIKYSVGSL--- 358
```

```
RESULT 13
S32763
kinectin 1 - human
A;Species: Homo sapiens (man)
A;Cross-references: EMBL:Z22551
R;Futterer, A.; Kruppa, G.; Kramer, B.; Lemke, H.; Kronke, M.
Mol. Biol. Cell 6, 161-170, 1995
A;Title: Molecular cloning and characterization of human kinectin.
A;Reference number: I37947; MUID:95306853
R;Kruppa, G.; Fuetterer, A.; Lemke, H.; Kroenke, M.
submitted to the EMBL Data Library, April 1993
A;Description: Cloning and characterization of TAF, a novel transactivating protein.
A;Reference number: S32763
A;Accession: S32763
A;Status: preliminary
A;Molecule type: DNA
A;Accession: S32763
A;Cross-references: EMBL:Z22551; NID:g296163; PIDN:CAA80271.1; PID:g296164

Query Match             4.6%;  Score 161;  DB 2;  Length 1356;
Best Local Similarity  20.2%;  Pred. No. 0.3;
Matches  147;  Conservative  130;  Mismatches  279;  Indels  172;  Gaps  34;

Db  521  KFVAKE--NEVQSLHSKLTDTLVSKQQLEQRLMQLMESEQKRVNKEESLQMQVQDILEQN  578
Qy  359  -------------SPVSASVL----------------------KKTEQKEVELQHISQKDCQED  388
Db  579  EALKAQIQOFHSQIAAQTSASVLAEELHKVIAEKDKQIKQTEDSLASERDRLTSK--EEE  636
Qy  389  DTTVKEQNFLYK--ETQETEKKLLFISEPIPHPSNELRGLNEKMSNKCSMLSTAEDDIRQN  447
Db  637  LKDIQNMNFLLKAEVO-------KLQALANEQAAAAHEL----EKMQQS--VYVKDDKIRLL  685
Qy  448  FTQLPLHKNKQECILDISEHTLSENDLEELRV----DHYKCNIQASVH-----VSDFSTDNS  500
Db  686  EEQL------------QHEIS-NKMEEFKILNDQNKA-LKSEVQKLQTLVSEQPNKDV  729
Qy  501  GSQPK----QKSDTVLFPAKDLKEKDLHSIFTHDSGLLITINSSQEHLTVQAKAPFHTPPEE  557
Db  730  VEQMEKCIQEKDEKLKTVEELLETGLIQVATKEEELNAIRTENSSLT---------KE  778
Qy  558  PNECDFKNMDSLPSGKIHRKVKIILG------RNRKENLEPN----AEFDKRTEFITQE-  606
Db  779  VQDLKAKQNDQVSFASLVEELKKVIHEKDGKIKSVEELLEAELLKVANKEKTVQDLKQEI  838
Qy  607  -------------ENRICSSPVQSLLDLFQTSEEKSEFLGFTSYTEKSGICN----VL  647
Db  839  KALKEEIGNVQLEKAQQLSITSKVQELQNLLKGKEEQMNTMKAVLEEKEKDLANTGKWLQ  898
Qy  648  DIWEEENS  655
Db  899  DLQEENES  906

Qy  17  GIQVKNKNRPSLKLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQKDIKDL----G  72
Db  261 GIQ-------KSGTKLKTETDKENAEVK---FKDFLLSLKTMMFSEDEALCVVDLLKEKS  311

Qy  73  GRVEEFLSK---DISYLISNKKEAKFAQTLGRISPVPSPESAYTAETTSPHSHGSSF  128
Db  312 GVIQDALKKSSKGELTTLIHQLQEKD--KILAAVK----EDAAATKDRCKQLTQEMMTE  364

Qy  129 KSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYIEQKKELYL  188
```

```
RESULT 14
A43336
microtubule-vesicle linker CLIP-170 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C;Accession: A43336
R;Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A;Title: CLIP-170 links endocytic vesicles to microtubules.
A;Reference number: A43336; MUID:92405160
A;Accession: A43336
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1392 <PIE>
A;Cross-references: GB:M97501; NID:g180621; PIDN:AAA35693.1; PID:g180622

Query Match             4.5%;  Score 159.5;  DB 2;  Length 1392;
Best Local Similarity  17.3%;  Pred. No. 0.38;
Matches  127;  Conservative  137;  Mismatches  295;  Indels  175;  Gaps  28;

Db  365  KERSNVVMTRMKDRIGTLEKEHNVFQNKIHVSYQETGQOMQ-MKFQQVR---EQMEAEIAH  420
Qy  189  LKKSSTSVRDGCKRVGSGAOKTRTGRLKPFVKVEDMSOLYRPFYLQLTNMPFINYSIQK  248
Db  421  LKQENGILRDAVSNTTNQLESKQSAELNK-----LRQDYARLVNELTEK---TGKLQQ  470
Qy  249  PCSPFDVDKPSSMQKQTQVRLRIQTDGDKYGGTSIQLQLKEKKKGYCECCLQKYEDLET  308
Db  471  E----EVQKKNAEQAATOLKVQLQEAAERRWE---EVQSYIRKRTAEHEAAQQ--DLOS  519
Qy  309  HLLSEQHRNFAQSNQYQVVDDIVSK-------LVFDFVEYE-KQTPKKRIKYSVGSL---  358
Db  520  KFVAKE--NEVQSLHSKLTDTLVSKQQLEQRLMQLMESEQKRVNKEESLQMQVQDILEQN  577
Qy  359  -------------SPVSASVL----------------------KKTEQKEVELQHISQKDCQED  388
Db  578  EALKAQIQOFHSQIAAQTSASVLAEELHKVIAEKDKQIKQTEDSLASERDRLTSK--EEE  635
Qy  389  DTTVKEQNFLYK--ETQETEKKLLFISEPIPHPSNELRGLNEKMSNKCSMLSTAEDDIRQN  447
Db  636  LKDIQNMNFLLKAEVO-------KLQALANEQAAAAHEL----EKMQQS--VYVKDDKIRLL  684
Qy  448  FTQLPLHKNKQECILDISEHTLSENDLEELRV----DHYKCNIQASVH-----VSDFSTDNS  500
Db  685  EEQL------------QHEIS-NKMEEFKILNDQNKA-LKSEVQKLQTLVSEQPNKDV  728
Qy  501  GSQPK----QKSDTVLFPAKDLKEKDLHSIFTHDSGLLITINSSQEHLTVQAKAPFHTPPEE  557
Db  729  VEQMEKCIQEKDEKLKTVEELLETGLIQVATKEEELNAIRTENSSLT---------KE  777
Qy  558  PNECDFKNMDSLPSGKIHRKVKIILG------RNRKENLEPN----AEFDKRTEFITQE-  606
Db  778  VQDLKAKQNDQVSFASLVEELKKVIHEKDGKIKSVEELLEAELLKVANKEKTVQDLKQEI  837
Qy  607  -------------ENRICSSPVQSLLDLFQTSEEKSEFLGFTSYTEKSGICN----VL  647
Db  838  KALKEEIGNVQLEKAQQLSITSKVQELQNLLKGKEEQMNTMKAVLEEKEKDLANTGKWLQ  897
Qy  648  DIWEEENS  655
Db  898  DLQEENES  905

Qy  11  KGHFQGGIQVKN------EKNRPSLKLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQ  65
Db  735 KASSEGKSEMKKLRQQLEAAEKQIKHLEIEKNAESSKASSITRELQGRELKLTNLQENLS  794

Qy  66  KDIKDLG-GRVEEFLSKDISYLISNKEAKFAQTLGRISPVPSPESAYTA-----ETTSP-  119
Db  795 E----VSQVKETLEKELQTL-----KERFAE-------ASEEAVSVQRSMQETVNKL  835
```

```
Qy  120 HPSHDGSSFKSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYI 179
        :::  ::: :  : :: :
Db  836 HQKEEQFNMLSSDLEKLRENLADMEAKFREKD----------------------------- 867

Qy  180 EQKKKELYLLKKSSTSVRDGGKRVG-SGAQKTRTG------RLKKPFVKVEDMSQLYRPFYLQ 235
        :: :  : :
Db  868 EREEQLIKAKEKLENDIAEIMKMSGDNSSQLTKMNDELRLKE----QLK 918

Qy  236 LT-----NMPFINYSIQKPCSPFDVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKK 291
        ::                                               : :
Db  919 LTKANENASFLQKSIEDMTVKAEQSQQEAAKKHEEEKKELER------KLSDLEKKMET 971

Qy  292 KKGYCECCLQKYEDLETHLLSEQHRNFAQSNQYQVVD-----------DIVSKLVFDFVEYE 342
         :   : :::  : : :
Db  972 SHNQCQELKARYE-RATSETKTKHEEILQNLQKTLLDTEDKLKGAREENSGLLQELEELR 1030

Qy  343 KDTPKKKRIKYSVGSLSPVSASVLKKTEQKEKVELQHISQKDCQEDDTTVKEQNFLYKET 402
        : : :: :          :: ::: ::
Db 1031 KQAEKAKAAQTAEDAMQIMEQMTKEKTETLASLEDTKQTNAKLQNELDTLKENNLKNVEE 1090

Qy  403 QETEKKLLFI-SEPIPHPSNELRGLNEKMSNKCSMLSTAE----------DDIRQNFTQLPLH 454
         : : :::                                         :::
Db 1091 LNKSKELLTVENQKMEEFRKEIETLKQAAAQKSQQLSALQEENVKLAEELGRSRDEVTSH 1150

Qy  455 KNKQECILDISEHTLSENDLEELR--------------VDHYKCNIQASVHV-SDFSTDNSGSQP 504
         :: : :   ::  :
Db 1151 QKLEE-----ERSVLNNQLLEMKKRESKFIKDADEEKASLQKSISITSALLTEKDAELE 1204

Qy  505 KQKSDTVLFPAKDLKEKDLHSIF-THDSGLITINSSQEHLTVQAKAPFHTPEEPNECDF 563
         :   :                  :::
Db 1205 KLRNEVTVLRGENASAKSLHSVVQTLESDKVKLELKVKNLELQLK 1251

Qy  564 KNMDSLPSGKIHRKVKIILGRNRKENLEPNAEFDKRT-----EFIT-----QEENRICS 612
         :   : : :  :                                 :  :
Db 1252 KRQLSSSSG-----NTDTQADEDERAQESQIDFLNSVIVDLQRKNQDLK 1295

Qy  613 SPVQSL-------LDLFQTSEEKSEFLGFTSYTEKSGICNVLDIWEEENSDNLLT- 660
         :  :        :  : :
Db 1296 MKVEMMSEAALNGNGDDLNNYDSDDQEKQ-----SKKKPRLFCDICDCFDLHDTEDCPTQ 1350

Qy  661 -AFFSSPSTSTFTG 673
          :: ::   : :
Db 1351 AQMSEDPPHSTHHG 1364
```

RESULT 15
S22695
restin - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999
C;Accession: S22695; S19853
R;Bilbe, G.; Delable, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A;Title: Restin: a novel intermediate filament-associated protein highly expressed in th
A;Reference number: S22695; MUID:92289675
A;Accession: S22695
A;Molecule type: mRNA
A;Residues: 1-1427 <BIL>
A;Cross-references: EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999
C;Keywords: cytoskeleton

```
Query Match         4.5%;  Score 159.5; DB 2;  Length 1427;
Best Local Similarity   17.3%;  Pred. No. 0.39;
Matches 127;  Conservative 137;  Mismatches 295;  Indels 175;  Gaps 28;

Qy  11 KGHFQGGIQVKN--------EKNRPSLKSLKTDNRPEKSKCKPLWGKVFYLDLPSVTISEKLQ 65
        : :                 :::
Db 770 KASSEGRSEMKLRQQLEAAAEKQIHLEIEKNAESSKASSITRELQGRELKLTNLQENLS 829

Qy  66 KDIKDLGGRVEEFLSKDIDSYLISNKKEANKFAQTLGRISPVPSPESAYTA------ 119
        : :::                                            :  :
Db 830 E---------VSQVKETLEKELQIL------KEKFAE 870
```

```
Qy  120 HPSHDGSSFKSPDTVCLSRGKLLVEKAIKDHDFIPSNSILSNALSWGVKILHIDDIRYI 179
        :::  ::: :  : :: :
Db  871 HQKEEQFNMLSSDLEKLRENLADMEAKFREKD----------------------------- 902

Qy  180 EQKKKELYLLKKSSTSVRDGGKRVG-SGAQKTRTG------RLKKPFVKVEDMSQLYRPFYLQ 235
        :: :  : :
Db  903 EREEQLIKAKEKLENDIAEIMKMSGDNSSQLTKMNDELRLKE----QLK 953

Qy  236 LT-----NMPFINYSIQKPCSPFDVDKPSSMQKQTQVKLRIQTDGDKYGGTSIQLQLKEKK 291
        ::                                               : :
Db  954 LTKANENASFLQKSIEDMTVKAEQSQQEAAKKHEEEKKELER------KLSDLEKKMET 1006

Qy  292 KKGYCECCLQKYEDLETHLLSEQHRNFAQSNQYQVVD-----------DIVSKLVFDFVEYE 342
         :   : :::  : : :
Db 1007 SHNQCQELKARYE-RATSETKTKHEEILQNLQKTLLDTEDKLKGAREENSGLLQELEELR 1065

Qy  343 KDTPKKKRIKYSVGSLSPVSASVLKKTEQKEKVELQHISQKDCQEDDTTVKEQNFLYKET 402
        : : :: :          :: ::: ::
Db 1066 KQADKAKAAQTAEDAMQIMEQMTKEKTETLASLEDTKQTNAKLQNELDTLKENNLKNVEE 1125

Qy  403 QETEKKLLFI-SEPIPHPSNELRGLNEKMSNKCSMLSTAE----------DDIRQNFTQLPLH 454
         : : :::                                         :::
Db 1126 LNKSKELLTVENQKMEEFRKEIETLKQAAAQKSQQLSALQEENVKLAEELGRSRDEVTSH 1185

Qy  455 KNKQECILDISEHTLSENDLEELR--------------VDHYKCNIQASVHV-SDFSTDNSGSQP 504
         :: : :   ::  :
Db 1186 QKLEE-----ERSVLNNQLLEMKKRESKFIKDADEEKASLQKSISITSALLTEKDAELE 1239

Qy  505 KQKSDTVLFPAKDLKEKDLHSIF-THDSGLITINSSQEHLTVQAKAPFHTPEEPNECDF 563
         :   :                  :::
Db 1240 KLRNEVTVLRGENASAKSLHSVVQTLESDKVKLELKVKNLELQLK 1286

Qy  564 KNMDSLPSGKIHRKVKIILGRNRKENLEPNAEFDKRT-----EFIT-----QEENRICS 612
         :   : : :  :                                 :  :
Db 1287 KRQLSSSSG-----NTDTQADEDERAQESQIDFLNSVIVDLQRKNQDLK 1330

Qy  613 SPVQSL-------LDLFQTSEEKSEFLGFTSYTEKSGICNVLDIWEEENSDNLLT- 660
         :  :        :  : :
Db 1331 MKVEMMSEAALNGNGDDLNNYDSDDQEKQ-----SKKKPRLFCDICDCFDLHDTEDCPTQ 1385

Qy  661 -AFFSSPSTSTFTG 673
          :: ::   : :
Db 1386 AQMSEDPPHSTHHG 1399
```

Search completed: December 27, 2001, 16:54:43
Job time: 273 sec

us-09-830-647-1.rpr